Loop-Erased Self-Avoiding Random Walks in Four and Five Dimensions

R. E. BRADLEY*
Department of Mathematics & Computer Science, Adelphi University, Garden City, New York 11530

S. WINDWER
Department of Chemistry, Adelphi University, Garden City, New York 11530

Received 28 September 1995; accepted 29 January 1996

ABSTRACT

Monte Carlo simulations of loop-erased self-avoiding random walks in four and five dimensions were performed, using two distinct algorithms. We find consistency between these methods in their estimates of critical exponents. The upper critical dimension for this phenomenon is four, and it has been shown that the mean square end-to-end distance grows as $n(\log n)^{\alpha}$. It has recently been established that the mean square end-to-end distance is asymptotically bounded by $n(\log n)^{1/3}$ (see Ref. 21). Our results show that asymptotic convergence to $n(\log n)^{1/3}$ in fact obtains and does so rather quickly. In five dimensions we examine the rate of asymptotic convergence to the mean-field model. © 1996 by John Wiley & Sons, Inc.

Introduction

The numerical simulation of the self-avoiding walk (SAW) on a lattice has been the subject of intense study for many years. This popularity may be attributed to the fact that the SAW provides a useful, simple model for a variety of complex phenomena: in chemistry, the flexible polymer molecule[1,2]; in physics, the $N \to 0$ limit of the N-vector model[3-5]; and, in mathematics, a concrete example of a non-Markovian process.[6]

* Author to whom all correspondence should be addressed.

It is widely believed that the growth rate of a SAW follows a power law relation with a critical dimension of four. More precisely, consider the usual SAW[7] on the d-dimensional integer lattice. For a walk of m steps, we let R_m denote the end-to-end distance; that is, the Euclidean distance from the initial point to the final point of the walk. If $\langle R_m^2 \rangle$ denotes the mean square end-to-end distance averaged over all SAWs of length m, then it is generally agreed that:

$$\langle R_m^2 \rangle \sim Cm^{2\nu} \qquad \text{if } d \le 4 \tag{1}$$

$$\langle R_m^2 \rangle \sim Cm(\log m)^{\alpha} \qquad \text{if } d = 4 \tag{2}$$

CCC 0192-8651 / 96 / 151750-07

and

$$\langle R_m^2 \rangle \sim Cm \qquad \text{if } d > 4 \qquad (3)$$

All constants are positive and independent of m. C depends on the local geometry of the lattice and ν depends only on dimension. We write $f(m) \sim g(m)$ to mean that $\lim_{m \to \infty} f(m)/g(m)$ exists and is equal to 1.

Clearly, $2\nu = 2$ for $d = 1$. For $d = 2$, Nienhuis[8,9] has shown that $2\nu = 3/2$, in agreement with the heuristic result of Flory[1] that $2\nu = 6/(d+2)$ for $1 \le d \le 4$ and $2\nu = 1$ when $d \ge 5$. Flory's argument is not valid in all dimensions, but its results appear to be correct in all cases except when $d = 3$ and 4. In the case $d = 3$, the best Monte Carlo work,[10] the best series solutions,[11,12] and the best renormalization group calculations[13] give $2\nu = 1.180 \pm 0.008$. In the case $d = 4$, the logarithmic correction noted above is needed. It is generally accepted, in analogy with other statistical models, that $\alpha = 1/4$, although recent work[14] suggests that α may be much closer to 1/3. In the case $d \ge 5$, Hara and Slade[15,16] have recently given a rigorous proof that Flory's conjecture is correct in these high dimensions.

Lawler[17] introduced another type of self-avoiding walk called the loop-erased self-avoiding walk (LESAW). To generate a LESAW, one performs a simple random walk (SRW) of n steps, which we call a *generating* random walk, and scans the list of nodes visited, searching for self-intersections. Wherever a self-intersection occurs, the intervening "loop" is erased. After all such loops have been erased, including immediate reversals, the result is a self-avoiding walk which we call the *derived* SAW. We denote the number of steps in this derived walk by N_n.

The growth of LESAWs is qualitatively similar to that of SAWs, but the exponents in the asymptotic relations (1) and (2) are larger. Lawler gives a heuristic account of why the terminal point of a LESAW should be driven away from the origin faster than that of a SAW (Ref. 18, pp. 175–178).

It is easy to see that relation (1) holds for LESAWs when $d = 1$ with $2\nu = 2$. Furthermore, Lawler[18] proves that relation (3) holds and demonstrates that in dimensions two and three $\langle R_m^2 \rangle$ grows at least as fast as the Flory estimates for these dimensions. Lawler conjectures that the growth rate is somewhat faster, and Monte Carlo estimates[19,20] suggest that relation (1) holds with 2ν being approximately 1.6 in two dimensions and 1.23 in three dimensions.

In the case of relation (2), Lawler[21] has recently shown that $\langle R_m^2 \rangle$ is asymptotically bounded by $m(\log m)^{1/3}$. More precisely, he has shown that there are positive constants c_1 and c_2 such that $c_1 m(\log m)^{1/3} \le \langle R_m^2 \rangle \le c_2 m(\log m)^{1/3}$. He conjectures that $n^{-1}(\log n)^{-1/3}\langle R_m^2 \rangle$ actually converges to a limit.

The first goal of this article is to examine the behavior of the four-dimensional LESAW and to look at its convergence. We perform Monte Carlo experiments using two distinct loop-erasing algorithms: batch and dynamic erasure (see next section). Additionally, we perform similar experiments in five dimensions both as a further comparison of the two algorithms and as a way to determine how quickly the asymptotic growth rate asserts itself and to approximate the value of C in relation (3).

Dynamic and Batch Erasure

Suppose one wishes to examine the validity of the relations (1) or (3) for SAWs or LESAWs. One approach would be to generate a sample of walks of length m and estimate $\langle R_m^2 \rangle$ with the sample average of the square end-to-end distance. Repeating this procedure for various values of m would yield data to estimate C and ν by a log–log plot. For the fixed-length methods used to generate SAWs (see Ref. 10), this is unproblematic.

The situation for LESAWs is different. Because the length N_n of the SAW derived from a classical walk of a length n is a random variable, one cannot predict how long to run the generating SRW in order that the derived SAW reach a prescribed length m. Instead, the programmer may generate data by running a classical walk, examining each node as it is generated to check for loops, which are erased as they occur. One continues this way, monitoring the evolving SAW, until it reaches the prescribed length m (which is N_n for some unpredictable n). In Ref. 19 we called this algorithm *dynamic loop erasing*.

A different approach, followed in Ref. 20, is to generate samples of classical random walks of various lengths n and use the data to estimate the growth of N_n as a function of n. We call this algorithm *batch loop erasing*.

To see the connection between these algorithms, suppose that relation (1) holds in some dimension $d \ne 4$. Then:

$$\langle R_m^2 \rangle_{\text{LESAW}} \sim Cm^{2\nu}$$

Thus:

$$m \sim C' \langle R_m^2 \rangle_{\mathrm{LESAW}}^{1/2\nu}$$

or:

$$\langle N_n \rangle \sim C' \langle R^2_{\langle N_n \rangle} \rangle_{\mathrm{LESAW}}^{1/2\nu} \quad (4)$$

On the other hand, it is well known from the theory of simple random walks that:

$$\langle R_n^2 \rangle_{SRW} = n \quad (5)$$

(More generally, $\langle R_n^2 \rangle_{SRW} = \sigma^2 n$, where σ^2 is the variance associated with each step of the random walk, but $\sigma^2 = 1$ in the orthogonal lattices which we are considering.)

$R^2_{N_n}$ for a derived LESAW has the same values as R_n^2 for the n-step classical random walk which generated it; this provides a qualitative link between the two algorithms and suggests that eq. (5) simplifies relation (4) as follows:

$$\langle N_n \rangle \sim C' n^{1/2\nu} \quad (6)$$

A similar argument may be made in the case $d = 4$, complicated only slightly by the logarithmic correction in model (2).

The dynamic-erased and batch-erased algorithms have been compared for dimensions 2 and 3 in Ref. 20, in which dynamic results were generated and compared to batch results in Ref. 21. We found close agreement for $d = 3$, but in two dimensions the dynamic algorithm gives a somewhat smaller value of ν than the batch algorithm.

Experiment and Results

Using the dynamic algorithm we generated 100,000 LESAWs of 5000 steps each in both four and five dimensions. Once the derived walk had reached 5000 steps, the squared distance R_m^2 was measured for a variety of values of m. These measurements were accumulated over the entire set of 100,000 walks. The means and standard deviations of R_m^2 are given in Tables I and II for dimensions four and five, respectively.

We also generated data using the batch algorithm in dimensions four and five. The results are given in Tables III and IV, respectively. In each dimension, we generated batches of 100,000 SRWs of length n = 400, 800, 1600, 3200, 6400, and 128,000 steps each. Loops were erased and the mean values of N_n were calculated, along with

TABLE I.
Mean Square End-to-End Distances for Dynamic Loop-Erased LESAWs in Four Dimensions.

m	$\langle R_m^2 \rangle$	SD
10	14.60	8.065
20	31.59	18.74
40	67.63	41.83
50	86.30	53.85
80	144.0	91.18
100	182.6	117.1
160	302.8	195.5
200	383.9	249.8
300	591.3	389.2
400	802.6	529.5
500	1016	672.8
600	1234	820.8
700	1449	963.5
800	1667	1112
900	1890	1260
1000	2113	1414
1100	2334	1567
1200	2553	1709
1300	2779	1863
1400	3005	2011
1500	3232	2160
1600	3460	2319
1700	3685	2472
1800	3911	2623
1900	4138	2771
2000	4369	2928
2100	4591	3074
2200	4818	3228
2300	5045	3377
2400	5277	3530
2500	5508	3686
2600	5744	3849
2700	5978	4027
2800	6208	4171
2900	6442	4334
3000	6675	4495
3100	6902	4653
3200	7133	4813
3300	7366	4970
3400	7598	5123
3500	7829	5277
3600	8063	5444
3700	8308	5613
3800	8547	5779
3900	8778	5947
4000	9011	6119
4100	9249	6280
4200	9473	6434
4300	9691	6585
4400	9916	6746
4500	10,140	6907
4600	10,360	7055
4700	10,580	7201
4800	10,800	7360
4900	11,010	7514
5000	11,210	7651

SD = standard deviation.

TABLE II.
Mean Square End-to-End Distances for Dynamic Loop-Erased LESAWs in Five Dimensions.

m	$\langle R_m^2 \rangle$	SD
10	13.04	6.807
20	27.21	15.28
40	55.91	32.56
50	70.30	41.39
80	114.0	68.19
100	143.4	86.26
160	232.2	141.4
200	291.8	178.7
300	440.0	272.5
400	588.1	363.7
500	735.9	456.7
600	884.8	551.7
700	1035	644.6
800	1185	736.1
900	1335	830.9
1000	1486	924.2
1100	1637	1018
1200	1786	1114
1300	1938	1210
1400	2088	1304
1500	2240	1402
1600	2388	1493
1700	2540	1591
1800	2691	1685
1900	2845	1786
2000	2995	1878
2100	3146	1970
2200	3296	2064
2300	3447	2159
2400	3591	2247
2500	3742	2341
2600	3892	2434
2700	4044	2530
2800	4192	2626
2900	4340	2721
3000	4490	2815
3100	4638	2905
3200	4786	2999
3300	4932	3096
3400	5087	3192
3500	5238	3289
3600	5387	3375
3700	5531	3462
3800	5682	3557
3900	5832	3654
4000	5980	3751
4100	6127	3847
4200	6278	3942
4300	6422	4035
4400	6569	4123
4500	6715	4211
4600	6865	4302
4700	7015	4404
4800	7168	4495
4900	7325	4594
5000	7470	4690

SD = standard deviation.

TABLE III.
Mean Number of Steps After Loop Erasure for Batch Loop-Erased LESAWs in Four Dimensions.

n	$\langle N_n \rangle$	SD
400	213.5	43.07
800	408.9	77.58
1600	786.3	140.7
3200	1516	258.0
6400	2932	475.4
12,800	5692	876.2

SD = standard deviation.

their standard deviations. Unlike the experiments performed in Ref. 21, a separate run was performed for each value of n. We hope this will maximize the likelihood of the independence of the measurements.

To test how well the data in Table I fits model (2), we began with the assumption that:

$$\langle R_m^2 \rangle = C(\log m)^{\alpha} m^{\beta} \tag{7}$$

Taking logs of both sides, this becomes:

$$\log\langle R_m^2 \rangle = K + \alpha \log\log m + \beta \log m$$

To estimate the exponents α and β, therefore, we ran a multiple linear regression with $\log m$ and $\log\log m$ as the independent variables and $\log\langle R_m^2 \rangle$ as the dependent one. The fit was extremely good, with the coefficient of determination (R^2) computed as greater than 0.999995.

The parameters α and β were estimated to be 0.3274 ± 0.0093 and 1.0032 ± 0.0017, respectively. With this estimate on β, we accept the hypothesis that model (2) is valid. Furthermore, our estimate on α is in accord with Lawler's proof that $\alpha = 1/3$.

Having empirically confirmed that $\beta = 1$, we fit the data in Table I to model (2). To accomplish this, we substitute $\beta = 1$ in (7) and solve:

$$\log\langle R_m^2 \rangle - \log m = K + \alpha \log\log m \tag{8}$$

A linear regression was performed, with $\log\log m$ as the independent variable and $\log\langle R_m^2 \rangle - \log m$ as the dependent one. Using all of the data in Table I gives $\alpha = 0.3446 \pm 0.0016$. The fit is even closer to $1/3$ if one discards the smaller values of m in hopes of better estimating asymptotic behavior. If we fit the data points $500 \leq m \leq 5000$ into the regression equation [eq. (8)], we get $\alpha = 0.3376 \pm 0.0051$; any reasonable confidence interval based on this estimate includes the value $1/3$.

TABLE IV.
Mean Number of Steps After Loop Erasure for Batch Loop-Erased LESAWs in Five Dimensions.

n	$\langle N_n \rangle$	SD
400	275.2	30.65
800	545.0	51.18
1600	1082	85.61
3200	2153	144.6
6400	4291	238.3
12,800	8557	401.3

We fitted the batch data in Table III to the model:

$$\langle N_n \rangle \sim Cn(\log n)^{\gamma}$$

Taking logs of both sides, this amounts to fitting the data to the model:

$$\log\langle N_n \rangle - \log n = K + \gamma \log\log n$$

If the argument which led to relation (6) is correct then we expect $\gamma = -\alpha$. The least squares estimate for γ is -0.3437 ± 0.0025. The value $-1/3$ falls within a 99% confidence interval based on these data.

In the case $d = 5$, we first fit the data to model (1) in hopes of verifying that $2\nu = 1$. Taking logs, the model becomes:

$$\log\langle R_m^2 \rangle = K + 2\nu \log m.$$

We performed a weighted linear regression of $\log\langle R_m^2 \rangle$ against $\log m$, weighting with the estimated variance of $\log\langle R_m^2 \rangle$. We used the propagation of error method and estimated the variance as:

$$\sigma^2_{\log\langle R_m^2 \rangle} \approx \frac{\sigma^2_{\langle R_m^2 \rangle}}{\langle R_m^2 \rangle^2}$$

Using this technique, we estimated 2ν based on all of the data in Table II. The result—1.0165 with a standard error of 0.001— is not particularly close to the rigorously determined theoretical value. Bearing in mind that model (3) concerns asymptotic behavior only, we repeated the linear regression discarding some of the data corresponding to small values of m. The results of these various fittings are summarized in Table V. Using the data for $1000 \leq m \leq 5000$ gave a particularly good fit, with the theoretical value of 2ν falling within 3 standard errors of the experimental value of 1.0015.

TABLE V.
Asymptotic Behavior of Dynamic-Erased LESAWs in Five Dimensions.

Data used	Estimated 2ν	Estimated C
All	1.0165 ± 0.0010	1.4960 ± 0.0005
100–5000	1.0087 ± 0.0006	1.4960 ± 0.0006
500–5000	1.0050 ± 0.0006	1.4956 ± 0.0007
1000–5000	1.0015 ± 0.0005	1.4939 ± 0.0008

For each of the data sets mentioned in Table V, we estimated the constant C in model (3) by performing a simple linear regression of $\langle R_m^2 \rangle$ versus m. We find that C is approximately 1.5.

Finally, we examined the data in Table IV to compare the batch-erased results for five dimensions to the dynamic-erased data. We fit the data to the model:

$$\langle N_n \rangle \sim Cn^{\mu}$$

using weighted least squares and estimated $\mu = 0.9928 \pm 0.0009$. Thus, $1/\mu$ falls somewhere between the estimates for 2ν, as given in lines 2 and 3 of Table V, just as one would hope. Performing a simple linear regression on $\langle N_n \rangle$ versus n gives a slope of 0.66774 ± 0.00038, also in very good agreement with the results tabulated in Table V.

Corrections to Scaling

Given that relation (2) for $d = 4$ and relation (3) for $d = 5$ are asymptotic, what does one do with small values of m? How large must a particular m be before it can be considered "sufficiently large"? In the previous section, we dealt with these issues in a simple *ad hoc* manner, ignoring the smaller values and examining the change in the exponent as a function of threshold.

It is possible to be much more systematic: renormalization group theory predicts a correction term of the form $\log|\log m|/\log m$ in four dimensions[22] and one of the form $m^{-1/2}$ in five. Both of these terms go to zero as m tends to infinity, the former very slowly. Our final statistical tests, then, involve fitting the full data sets to these models.

Under the assumption that:

$$\langle R_m^2 \rangle = C_1 m(\log m)^{1/3}\left(1 + C_2 \frac{\log|\log m|}{\log m}\right)$$

in four dimensions, we may estimate the constants C_1 and C_2 by performing a simple linear regression with:

$$\frac{\log|\log m|}{\log m}$$

as the independent variable and:

$$\frac{\langle R_m^2 \rangle}{m(\log m)^{1/3}}$$

as the dependent variable. We find that $C_1 = 1.1414 \pm 0.0041$ and $C_2 = -0.1066 \pm 0.0134$.

To complete our investigation of the LESAW in four dimensions, we repeated the multiple linear regression using eq. (7), except that the independent variable was adjusted as follows:

$$\frac{\langle R_m^2 \rangle}{1 - \dfrac{0.1066 \log|\log m|}{\log m}}.$$

We found an exceedingly good fit, with the coefficient of determination in excess of 0.999995. We calculated $\alpha = 0.3281 \pm 0.0088$ and $\beta = 1.0008 \pm 0.0016$, which is wholly consistent with $\alpha = 1/3$ and $\beta = 1$.

In five dimensions, correction to scaling suggests that the data, including small values of m, should fit the model:

$$\langle R_m^2 \rangle = C_1 m\left(1 + \frac{C_2}{\sqrt{m}}\right)$$
$$= C_1 m + C'\sqrt{m}$$

We performed a multiple linear regression using m and $\sqrt{m}$ as the independent variables and $\langle R_m^2 \rangle$. Forcing the regression equation through the origin, we find $C_1 = 1.4970 \pm 0.0015$ and $C_2 = -0.0942 \pm 0.0595$, with a coefficient of determination exceeding 0.999995. Allowing for a constant term, we calculate $C_1 = 1.4906 \pm 0.0021$ and $C_2 = 0.2912 \pm 0.1120$. In this setting, the constant term is -11.25 ± 2.90 and the coefficient of determination is 0.99999. In both cases, the fit is extremely good and the value of C in model (3) is approximately 1.5.

Conclusions

We find that the dynamic-erased and batch-erased algorithms are in agreement when estimating the critical exponents for LESAWs in four and five dimensions. In Ref. 20 we found that this was also the case in three dimensions but not in two dimensions. It is quite likely that this disagreement, when $d = 2$, is a consequence of the fact that the simple random walk is recurrent in dimensions one and two, but transient in higher dimensions; this conjecture can be made with greater confidence now that we have verified that the phenomenon appears to be limited to the case $d = 2$.

We find compelling evidence to accept Lawler's conjecture that $\langle R_m^2 \rangle$ in four dimensions is not simply asymptotically bounded by $m(\log m)^{1/3}$, but is in fact asymptotically equal to $Cm(\log m)^{1/3}$. We find an extremely close fit to this function, with C taking a value of approximately 1.47.

We find that the behavior of the five-dimensional LESAW has settled reasonably well into its asymptotic limit by the time the walk has reached 1000 steps. We also find that such walks grow about 50% faster that the simple random walk in five dimensions, but we see no compelling reason to believe that C is exactly 3/2.

Acknowledgments

The authors are grateful to the second referee for his helpful suggestions. We also thank Professor A. D. Sokal for an enlightening discussion concerning corrections to scaling. We also acknowledge Henry Saltiel, director of the Mary Lou Buchanan Computing Centre, and his staff, for their continuing help and technical support.

References

1. P. J. Flory, *Statistical Mechanics of Chain Molecules*, Wiley, New York, 1969.
2. C. Domb, *Adv. Chem. Phys.*, **15**, 229 (1969).
3. P. G. de Gennes, *Phys. Lett.*, **38A**, 339 (1972).
4. P. G. de Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, NY, 1979.
5. J. des Cloizeaux, *J. Phys. (Paris)*, **38**, 281 (1975).
6. W. Feller, *An Introduction to Probability Theory*, Vol. I, 3rd ed., Wiley, New York, 1968.
7. S. Windwer, *Markov Chains and Monte Carlo Simulations in Polymer Science*, Marcel Dekker, New York, 1970, chapter 5.
8. B. Nienhuis, *Phys. Rev. Lett.*, **49**, 1062 (1982).
9. B. Nienhuis, *J. Stat. Phys.*, **34**, 731 (1984).
10. N. Madras and A. D. Sokal, *J. Stat. Phys.*, **50**, 109 (1988).

11. A. J. Guttman, *J. Phys. A: Math. Gen.*, **20**, 1839 (1987).
12. A. J. Guttman, *J. Phys. A: Math. Gen.*, **22**, 2807 (1989).
13. J. C. Le Guillou and J. Zinn-Justin, *Phys. Rev. B.*, **21**, 3976 (1980).
14. P. Grassberger, R. Hegger, and L. Schäfer, *J. Phys. A: Math. Gen.*, **27**, 7265 (1994).
15. T. Hara and G. Slade, *Commun. Math. Phys.*, **147**, 101 (1992).
16. T. Hara and G. Slade, *Rev. Math. Phys.*, **4**, 235 (1992).
17. G. F. Lawler, *Duke Math. J.*, **47**, 655 (1980).
18. G. F. Lawler, *Intersections of Random Walks*, Birkhäuser, Boston, 1991.
19. R. E. Bradley, S. Windwer, *Phys. Rev. E*, **51**, 241 (1995).
20. A. J. Guttmann and R. Bursill, *J. Stat. Phys.*, **59**, 1 (1990).
21. G. F. Lawler, *J. Fourier Anal. and Appl.*, special issue: Proceedings of the Conference in Honor of J.-P. Kahane, 347 (1995).
22. E. Brezin, J. C. Le Guillou, and J. Zinn-Justin, In *Phase Transitions and Critical Phenomena*, Vol. 6, C. Domb and M. S. Green, Eds., Academic Press, London, 1976, chapter 3.